

|  |  |
| --- | --- |
| **Proposal By** | **Proposal To** |
|  |  |
| |  | | --- | | **Name:** Swetha Venkatesappa (SX/EDS3) | | **Name:** Shashi Kumar R (BGSW/HRL-LCG1) |
| **Designation:** Project Manager | **Designation:** Project Manager |
| |  | | --- | | **Telephone:** +91-9663304554 | | **Telephone:** |
| |  | | --- | | **Mobile:** +91-9663304554 | | **Mobile:** |
| |  | | --- | | **E-Mail:** Venkatesappa.Swetha@in.bosch.com | | **E-Mail:** ShashiKumar.R@in.bosch.com |
| **Proposal Submitted on 26-Aug-2024** | |

**PROPOSAL DETAILS**

|  |  |
| --- | --- |
| Confidentiality: | Confidential |
| Copyright and reproduction: | Not allowed without the written consent from Bosch Global Software Technologies Ltd |
| Validity: | 90 days |
| Reference Number: | HRL-LCG1\_SX/EDS-MM\_23082024 |

Proposal to:

**JKL Learning**

Project for:

Fraud Detection in Reimbursement Invoice Submissions

Contents

[1. Executive Summary 4](#_Toc175309198)

[1.1 Company Profile: 4](#_Toc175309199)

[2. Project details 5](#_Toc175309200)

[2.1 Project Background 5](#_Toc175309201)

[2.2 Project Objective 5](#_Toc175309202)

[3. Project scope 6](#_Toc175309203)

[3.1 Scope: 6](#_Toc175309204)

[3.1.1 Detailed Scope 6](#_Toc175309205)

[3.2 Out of Scope: 7](#_Toc175309206)

[3.3 Architecture 8](#_Toc175309207)

[3.4 Pre-Requisites: 9](#_Toc175309208)

[3.5 Assumptions: 9](#_Toc175309209)

[3.6 Deliverables, Receivables and Acceptance Criteria 9](#_Toc175309210)

[3.7 Bill of Materials (BOM) 11](#_Toc175309211)

[4. Project schedule 12](#_Toc175309212)

[5. Project Management 13](#_Toc175309213)

[5.1 Change Request Procedure: 13](#_Toc175309214)

[5.3 Roles and Responsibilities 14](#_Toc175309215)

[5.4 Escalation Management 15](#_Toc175309216)

[5.5 Communication Plan 16](#_Toc175309217)

[5.6 Risks & Mitigations 17](#_Toc175309218)

[5.7 Information Security 17](#_Toc175309219)

[6 Commercials 18](#_Toc175309220)

[6.1 Price Summary 18](#_Toc175309221)

[6.2 Billing Milestones 18](#_Toc175309222)

[6.3 Commercial Terms and Conditions 19](#_Toc175309223)

[7 Case Studies 20](#_Toc175309224)

[8 Appendix 21](#_Toc175309225)

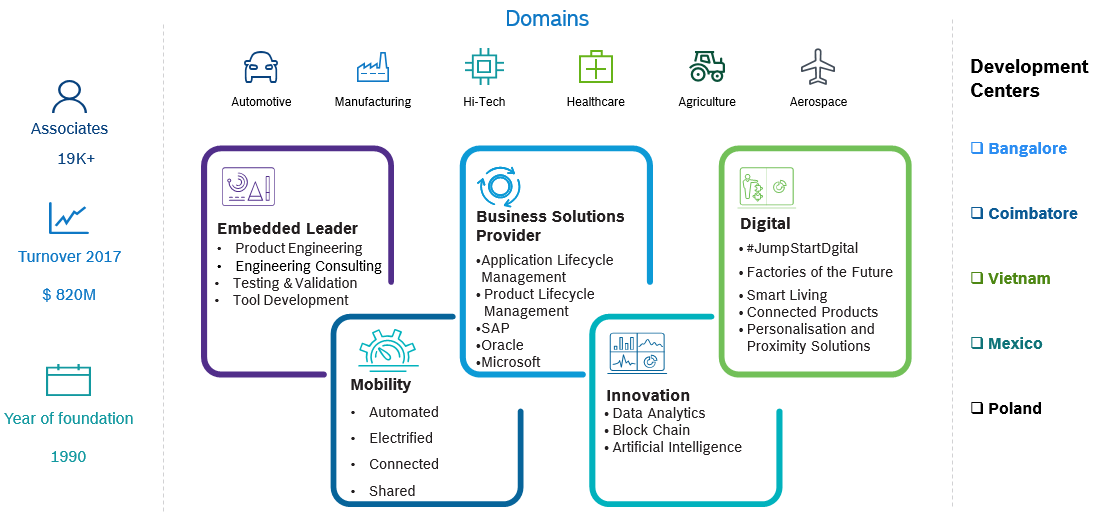
[9 Signatures 22](#_Toc175309226)

**GLOSSARY**

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No** | **Module** | **Abbreviation** | **Full form** |
| 1 | Generic | BGSW | Bosch Global Software Technologies |
| 2 | Generic | EDA | Exploratory Data Analysis |
| 3 | Generic | OS | Operating System |
| 4 | Generic | OCR | Optical Character Recognition |
| 5 | Generic | GST | Goods and Services Tax |
| 6 | Generic | SPOC | Single Point of Contact |
| 7 | Generic | SOW | Statement of work |
| 8 | Generic | AI | Artificial Intelligence |
| 9 | Generic | SME | Subject Matter Expert |
| 10 | Generic | KPI | Key performance Indicator |
| 11 | Generic | ML | Machine Learning |
| 12 | Generic | UAT | User Acceptance Testing |
| 13 | Generic | RFP | Request for Proposal |

# Executive Summary

## Company Profile:



# Project details

## 2.1 Project Background

Bosch offers employees the option to reimburse courses from various vendors, which involves handling extensive request and invoice data. Currently, the process of identifying frauds in the dataset is done manually. Manual identification of frauds can be time-consuming and may not capture all relevant patterns or anomalies in the data. The objective of the project is to build AI intelligence that can analyze the data and identify anomalous patterns in the request and invoice data and flag the potential fraudulent invoice filings. This accelerates the invoice review process and Bosch can potentially improve efficiency, reduce manual effort, and gain insights into areas where cost savings can be achieved.

## 2.2 Project Objective

The primary objective of the project is to develop an intelligent fraud detection algorithm for the identification of outliers/anomalies in request and invoice data.

# Project scope

## Scope:

Design and develop an intelligent fraud detection algorithm for the identification of anomalies in request and invoice data so that necessary actions can be triggered to mitigate the risks.

### 3.1.1 Detailed Scope

**Algorithm Development Scope**

Develop fraud detection algorithm to identify fraudulent activities based on Request data and Invoice data as the data source.

**Request Data based Fraud Detection:**

* Scope is to develop a request data-based fraud detection using three years of historic request data to identify anomalies for Udemy vendor.
* Scope includes implementing rule-based approaches for fraud detection, which includes below rules:

1. Duplicate invoice number used by employees.
2. Incorrect pattern in invoice number based on invoice date.

**OCR based Fraud Detection:**

* Scope involves developing OCR based fraud detection using invoice data for Udemy vendor.
* The key entities that need to be extracted from invoice includes invoice number, invoice date, recipient details, supplier details, course description, taxable value, GST amount and total amount.
* Scope includes implementing rule-based approaches for fraud detection, which includes below rules:

1. GST calculation errors in invoices.
2. Invoice numbers mismatched in invoice and request data.
3. Deviation from standard invoice templates.
4. Discrepancies between the amount in the invoice and the claimed amount in request data.
5. Mismatched font sizes in invoices using OCR libraries.
6. Mismatch between selected vendor and claimed vendor in invoice.

Data Ingestion includes:

* Setting up Connectivity from On-Prem MySql server to Azure Cloud via Integration Runtime.
* Data Ingestion: Ingestion of above source tables from MySQL DB to Azure cloud (Blob/ SQL)
* Copy PDF file from On-prem: Move PDF files to Azure cloud storage.
* Write Back: Update Flag and reasoning for invoices identified as Fraud to On-premise MySql
* Excel Report of processed invoices with status for each invoice will be updated in the common path shared.

Below Sources are considered:

|  |  |  |  |
| --- | --- | --- | --- |
| **S. No** | **Tables** | **On Prem Database** | **Mode** |
| 1 | JKL\_Source table | MySQL | Input/Output |
| 2 | JKL\_Cloud\_DB | Azure SQL | Output |

**General:**

1. All documentations and business communication will be in English.
2. Post Go-Live stabilization support for 2 weeks after agreed deployment date.

## Out of Scope:

1. Support co-ordination with OEM or 3rd party vendors for any product issues/bugs.
2. Source system data availability and data quality related issues.
3. Setup and management of any DevOps / integration tools.
4. Procurement and management of third-party licenses and certificates.
5. Processing of Data other than Udemy vendor invoices.
6. Processing of Data other than request and invoices data sources.
7. Detecting Invoice tamper activities which do not align with those covered in scope.
8. Any Module and functionality not described in the above scope section.
9. Any modifications / new developments to the existing source systems
10. Reporting:
    1. Mobile app / Report Development
    2. UI/UX development
    3. Web App Development
    4. Administration activities
11. SSO and AD integration
12. Custom Visual Development
13. Any Alerts, Alarms and Notifications
14. Realtime functionality in Power BI reports

**Model Deployment Scope**

* Trained algorithm to be deployed in customers Azure cloud tenancy.
* Deployed model to generate predictions daily once up to 1000 invoices.

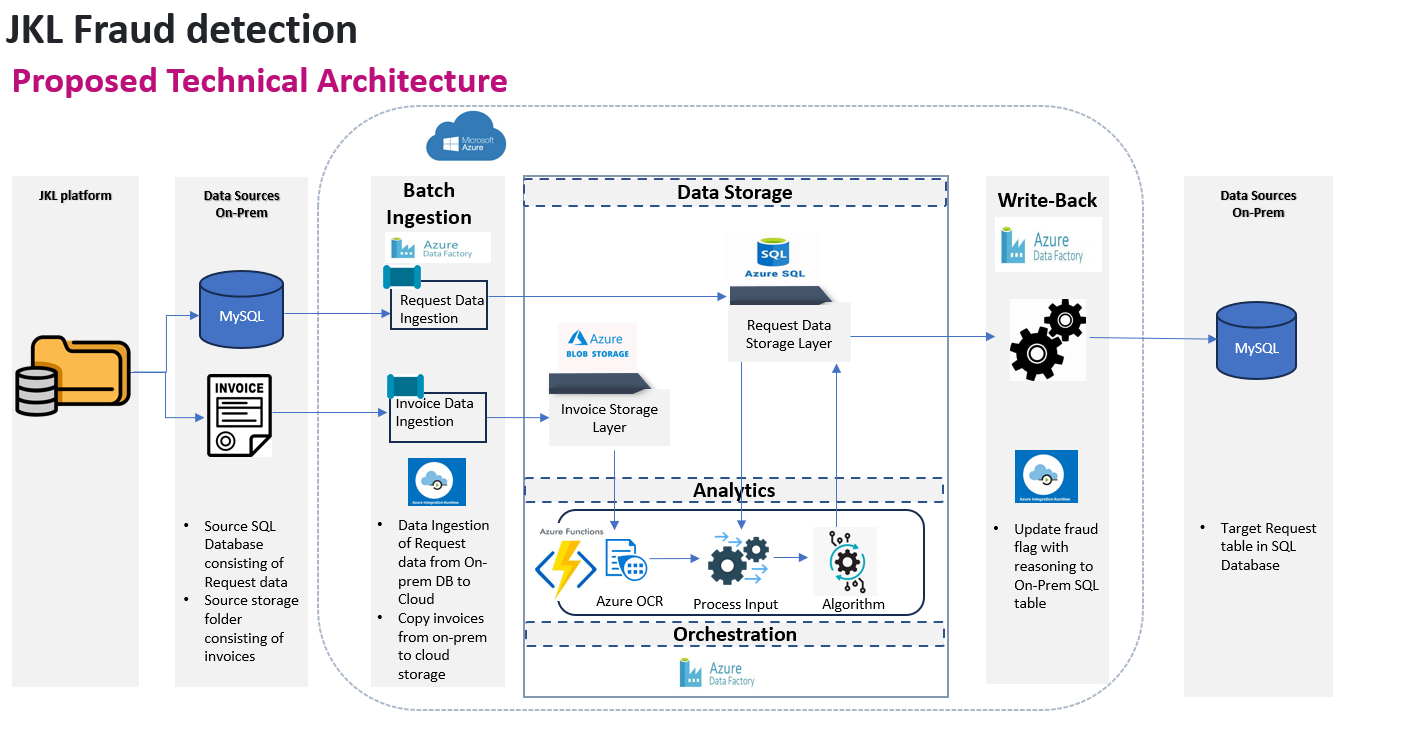
**Performance Metric**

The performance metric for the modules are as follows:

* The metrics for Fraud detection and OCR extraction algorithm are to be mutually agreed and documented in the requirement specification document post data quality assessment.

## Architecture

SX/EDS-MM proposes the below approach to build intelligent fraud detection system for JKL Learning:



|  |  |  |  |
| --- | --- | --- | --- |
| S.No | Component | Description | Tools/Lib |
| 1 | Data Ingestion and Write backup Module | Ingestion of above source tables from on Prem DB to cloud and Cloud to On-prem | Using Azure ADF |
| 2 | Data and file storage | Storing of Ingested data for Processing and PDF files | Azure cloud (Blob and SQL) |
| 3 | On-premise location access | Access On-premise location DB and PDF files | Azure runtime integration |
| 4. | ML Deployment | Hosts and manages ML models, OCR services, making them available for inference. | Azure Function |

## 

## Pre-Requisites:

Below are the key pre-requisites expected from JKL team:

|  |  |  |
| --- | --- | --- |
| **#** | **Project Phase** | **Pre-requisites** |
| 1 | Project Kick Off | 1. Single point of contact (SPOC) from JKL 2. Purchase order. 3. Access to source system artifacts 4. Three-year historical request and invoice data |
| 2 | Study | 1. Availability of Customer SPOC from customer for process understanding and data understanding 2. Agreement on Data quality and acceptance criteria on performance metric 3. Identification of Udemy invoices for request-based fraud detection. 4. Rules for implementing rule-based fraud detection to be provided by JKL. 5. Invoices should be named in well defined naming convention so that corresponding user id can be identified. 6. Access on premise sources Data and invoices 7. For all the type of Fraud detection rules, atleast 10 sample files to be shared for each rule. |
| 3 | Design | 1. Customer SPOC shall provide timely clarification of open points. |
| 4 | Development and Unit Testing | 1. Customer SPOC shall provide timely clarification of open points. 2. All the required access and permission to the production environment 3. Installation of required software libs 4. Access to On-prem Database/ Table 5. Access to Azure Subscription |
| 5 | UAT | 1. Availability of Key users 2. Daily upload of fresh data for model prediction. |
| 6 | Go-Live & Stabilization |  |
| 7 | Project Closure |  |

## Assumptions:

1. Scope creep will follow the CR process.
2. JKL team to ensure availability of all necessary infrastructure/Source Systems with necessary interfaces for data extraction as defined in the prerequisites.
3. Infrastructure to be procured from BGSW/BDO.
4. In case of scope creep, revised estimates for schedule and costs will be mutually discussed and agreed.
5. Customer is responsible for Data Ownership and Data Quality
6. Any product / feature limitations pertaining to OEM shall be routed to OEM for support.
7. Post data quality assessment acceptance criteria to be revisited.
8. Acceptance period in UAT and Production for deliverables is 5 calendar days. If no feedback received on deliverables submitted by BGSW within this acceptance period, the deliverable is deemed accepted.

## Deliverables, Receivables and Acceptance Criteria

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| # | Project Phase | Deliverables by BGSW | Receivables from JKL | Acceptance Criteria |
| 1 | Project Kick Off | * Project plan and schedule with milestones * Team structure | * Purchase Order * Project Level Agreement * Project Team Structure | * Deliverables are in line with the project scope. |
| 2 | Study | * Requirement Specifications Document (RSD) | * Clarifications on need basis from SMEs * Agreement on performance metrics * Sign off on Requirements Specification Document (RSD) * Three-year historical request and invoice data * Rules for implementing rule-based fraud detection. | * Solutioning covers identified requirements. |
| 3 | Design | High Level Design Document | * Clarifications on need basis | * Solutioning covers identified requirements. |
| 4 | Development and Unit Testing | * Data science model development * Pipeline Development  - Process Input, OCR, Model Processing, Storing Predictions * Unit Test cases derived from shared data and Test results | * Clarifications on need basis |  |
| 5 | UAT | * UAT support and issue resolution | * Signoff on UAT * Clarifications on need basis * Daily upload of fresh data for model | * No critical defects |
| 6 | Go-Live & Stabilization | * Deploy Solution in Production | * Signoff on Go-Live * Access to production environment * Certificates, Keys used in production phase. | * No critical defects |
| 7 | Project Closure | * Documentation * Transition to Operations | * Handover confirmation * Signoff on Project closure | * No deviation from the agreed scope |

## 3.7 Bill of Materials (BOM)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S. No | Hardware / Infrastructure | Environment | No’s of components | Specifications |
| 1 | Azure ADF, Azure SQL, Azure blob | Development | 3 | Total approximate cost is 50$ considering One GB of data and pipeline execution once in a day and region is East US |

# Project schedule

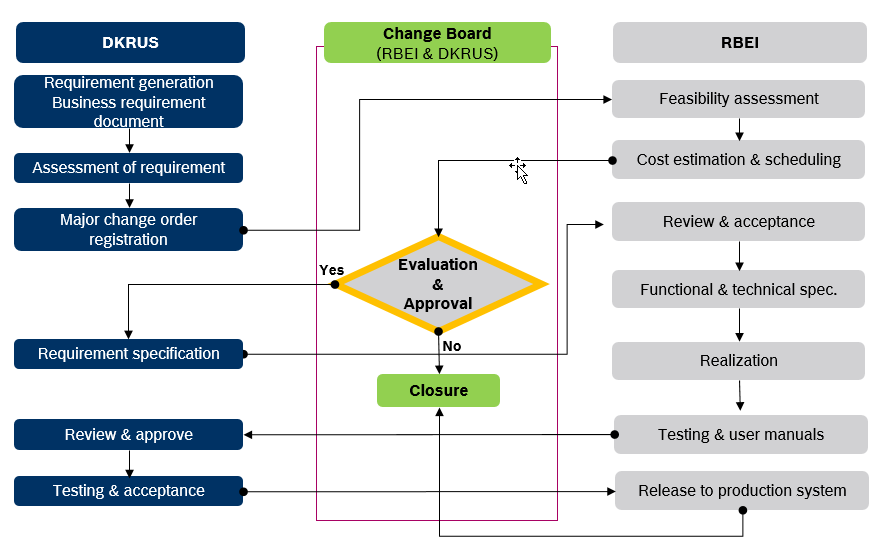
The indicative timeline for this project is as depicted below.



# Project Management

## 5.1 Change Request Procedure:

The Change Request procedure (shown below) applies to all changes resulting from requirements outside the scope defined in this proposal.



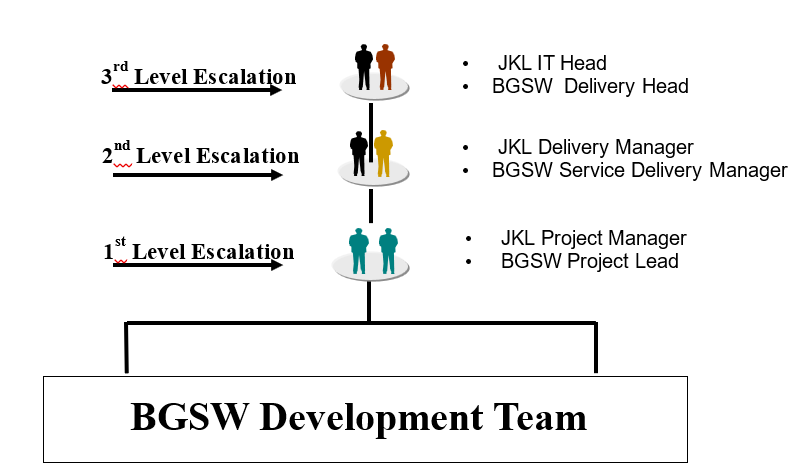
Any activity causing delay in project schedule will have an impact on project commercials and must be discussed mutually between BGSW and JKL. Additional efforts must be incorporated in the project through Change request process.

## Roles and Responsibilities

|  |  |
| --- | --- |
| **JKL** | |
| **Role** | **Responsibility** |
| **JKL Project Manager** | * Responsible for project delivery * Overall project responsibility * Participate in joint reviews with BGSW * Approves change requests |
| **JKL Core Team** | * Responsible for UAT * Reports defects to product owner during UAT |

|  |  |
| --- | --- |
| **BGSW** | |
| **Role** | **Responsibility** |
| **Project Lead** | * Responsible for overall project management * Administrative control of BGSW project team * Project inputs to JKL Program Manager * Provide resources for the project * Coordination with teams |
| **Architect** | * Designing solution (end-to-end vision for data flow & pipeline) * Development of data models for database structures * Integration with source systems * Integration of technical functionalities including scalability, security, reliability * Visualizations |
| **Data Engineer** | * Integration of Source Interfaces * Building/developing data processing and pipeline. * Implementation of the architected solution |
| **Data Scientist** | * Exploratory Data Analysis * Development of AI/ML models |
| **ML Engineer** | * Setting up of execution pipelines * Maintenance of cloud infrastructure in terms of availability during batch run |

## Escalation Management



## Communication Plan

An effective communication is necessary for successful execution and tracking of the project. The below communication plan is proposed for the project.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S No | Communication Type / Media | Participants | Frequency | Topics to be discussed |
| 1 | Project status review | JKL  BGSW | Weekly | Status of the project |
| 2 | Project team meeting | Project teams | Fortnightly | Activities / status / clarifications on project related activities |
| 3 | Project status reporting | JKL  BGSW | Monthly | Monthly status of the project |
| 4 | Issue escalation | JKL  BGSW | Ad-Hoc | Issues which are escalated |
| 5 | Steering committee | JKL business representative  JKL IT head  BGSW Delivery head  BGSW and JKL  project managers  Other key stakeholders | Quarterly | Decisions, approvals & status of the project |

## Risks & Mitigations

|  |  |
| --- | --- |
| Risks | Mitigation |
| Unavailability of necessary Infrastructure/Source Systems or the interfaces for extraction of necessary data from the source systems | JKL team to ensure availability of all necessary infrastructure/Source Systems with necessary interfaces for data extraction as defined in the prerequisites. Infrastructure to be procured from BGSW/BDO. |
| Missing clarity of Data field mapping and Calculation for KPIs | JKL team to ensure availability of all Data field mapping and calculations documented and made available to BGSW team before the Design phase |
| Delay in acceptance of deliverables resulting in schedule and effort deviations | JKL should adhere to the plan for review and sign-off. |
| Disturbance due to Natural calamities and social chaos | JKL to have necessary Business Continuity Plan. Schedule to be revisited, adjusted, and mutually agreed |

## Information Security

BGSW will adhere to the following information security practices.

* Will comply to JKL guidelines and procedures – Institutionalize measures by internal audits.
* Confidentiality: Signing of confidentiality letter by each associate

# Commercials

## 6.1 Price Summary

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Item** | **Topic** | **Duration** | **Total Cost** | **Remarks** |
| 1 | JKL Fraud Detection | 12 weeks | TBD |  |

## 6.2 Billing Milestones

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Item** | **Milestone** | **Timeline** | **Cost in INR per milestone** | **Remarks** |
| 1 | Project kick-off | T+0 | 35% | Advance payment with PO |
| 2 | UAT | T+ 7 weeks | 50% |  |
| 3 | Project Closure | T+12 weeks | 15% |  |

*Note: T indicates project kick-off date*

## Commercial Terms and Conditions

* Taxes as applicable are additional.
* In case the project is stopped for unforeseen circumstances, then invoice will be raised for the effort which has been spent.
* Any delay in the project schedule not attributable to BGSW will lead to a revision in the commercial estimates which will be discussed and agreed upon.
* Any change in scope or schedule from the proposal may involve re-evaluation of the pricing and revision (upward or downward) if required.
* Any change in the assumptions relating to responsibilities will involve Change Order processing and follow up proposal.
* Project start is subject to publishing of names of business users, project manager from JKL, sign off on schedule and single point of contact in IT.
* JKL shall bear cost of travel, accommodation and allowance in case any consultant is asked to travel to a location outside of the project location in JKL which is not planned as part of the project
* Any clarifications on the invoice to be completed within 10 days from the date of invoice.
* Invoices are payable without deduction within thirty (30) days from the date of invoice.

# Case Studies

# Appendix

# Signatures